

厚生労働行政推進調査事業費補助金（健康安全・危機管理対策総合推進事業）  
分担研究報告書

次世代シーケンサーに関わる各種検査法の開発  
ゲノム解析ワーキンググループ活動

研究分担者 岩本 朋忠 神戸市健康科学研究所

研究協力者 地方衛生研究所ゲノム解析ワーキンググループ

**研究要旨**

地方衛生研究所（地衛研）におけるゲノム解析の普及を目的として、ゲノム解析の1次処理にあたるウェットラボ操作に関するマニュアル作成に取り組んだ。地衛研全国6支部をカバーする検査実務者16名からなるゲノム解析ワーキンググループを構成し、各施設での解析体制に関する情報共有、ゲノムDNAの断片化処理条件やライブラリーのサイズセレクションについての実験的検証作業を経て、地衛研の体制に即したマニュアルが作成できた。

**A. 研究目的**

地方衛生研究所（地衛研）の体制に即したゲノム検査マニュアルの作成、ならびに、ゲノム解析手法に関する知見の共有と集積を図ることで、地衛研におけるゲノム解析の普及を推進する。

**B. 研究方法**

マニュアルの対象は、ゲノム解析の1次処理にあたるウェットラボ操作、すなわち、ゲノムDNAライブラリー作製と次世代シーケンサーでのリードデータ取得とした。地衛研全国6支部をカバーする検査実務者16名からなるゲノム解析ワーキンググループを構成し、マニュアル作成における課題の洗い出しと実験データの取得による検証作業を行った。検証作業は、次の2項目について大腸菌 ATCC25922 株を用いて行った。（検証作業項目1）

ゲノムDNA断片化におけるサーマルサイクラーのリード温度および反応時間の影響について

（検証作業項目2）

ライブラリーサイズセレクション方法の

違いによるリードデータへの影響について

（倫理面への配慮）

検査手法の標準化とマニュアル化についての研究であり、個人情報など倫理面での配慮を必要とする情報や行為は一切伴わない。

**C. 研究結果**

1. マニュアル作成における課題

マニュアル作成にあたっては、ライブラリーのサイズセレクションの方法、各ライブラリーの混合比率を均質化するためのノーマライゼーションのタイミング、ライブラリーの濃度測定方法について意思統一が必要であることが分かった。ワーキンググループメンバー間でのweb会議を繰り返し、サイズセレクションの方法については、ゲル切り出しによる方法とビーズ法の両方の手法を併記し、それぞれの特徴を記載することとした。ノーマライゼーションは、ライブラリー調整に使用するDNA濃度を均一化（20 ng/uL）することとした。濃度測定法は、Qubit Fluorometerを用いた方法に統

一し、定量的 PCR 法をトラブルシューティングのための方法として記載することとした。

2. ゲノム DNA 断片化におけるサーマルサイクラーのリッド温度および反応時間の影響について

ライブラリー作製のフラグメンテーションについては、公式マニュアルでは input DNA 量の違いにより 2~30 分という記載になっている。本マニュアルでは input DNA 量を 50 ng に固定しているため、反応時間が断片化に与える影響は大きい。そのため、大腸菌 ATCC25922 株を用いて、反応時間による断片化の進行具合を検証した。

図 1 の通り、断片化は反応時間 3 分でも十分進行しており、反応時間を 3 分に設定することの妥当性が確認できた。MiSeq を使用してリードデータを取得する場合、片側 250 または 300 bp となるペアエンドリードを想定するケースが多いため、少し長めの断片を残存させる条件で反応時間を設定することが推奨される。

続いて、Lid 温度が断片化に与える影響について検証した。同様に大腸菌 ATCC25922 株を用いてサーマルサイクラーのリッド（ふた）温度を 70℃、80℃、105℃に設定し反応を行った。105℃についてはフラグメンテーション反応中にリッドを開放した状態と閉めた状態の 2 条件で検証を行った。その結果を図 2 に示す。70℃、80℃、105℃（開放）では断片化の進行に大きな違いは見られなかったが、105℃で閉めた場合は若干断片化の進行が早くなる傾向が認められた。リッド温度を調節可能なサーマルサイクラーの場合は 80℃、調節できないサーマルサイクラーの場合は、32℃反応中、リッドを開放して、65℃のステップになってからリッドを閉めるという操作が推奨された。

3. ライブラリーサイズセレクション方法の違いによるリードデータへの影響

大腸菌 ATCC25922 株を用いて、ビーズ精

製とゲル抽出によるライブラリーの状態およびリードデータの質について検証した。両手法によるライブラリーサイズの分布の解析結果を図 3 に示す。ゲル抽出精製が 550bp 付近で比較的シャープな山を形成しているのに対し、ビーズ精製ではアダプターダイマーやプライマーダイマーなどは見られなかったものの、山がブロードで 1000 bp を超えるサイズのフラグメントも残存しているのがわかる。1000 bp を超えるフラグメントは Miseq/iSeq でのクラスター形成過程において Bridge PCR がうまくいかずクラスター形成不全を引き起こすものと思われる。そのため、想定したライブラリーの DNA 濃度から推定したクラスター密度と実際の密度の乖離が大きくなってしまう可能性がある。

続いて、作製したライブラリーから MiSeq Reagent Kit v3 (600 Cycles)を用いてリードデータを取得し、fastp v0.23 によるトリミングを実施した。その結果、リード数の残存率はトリミングの前後でどちらの手法も 98%となり大きな違いはなかったが、残存した総リード長はゲル抽出が約 94%であるのに対し、ビーズ精製は約 74%しか残存せず、大きく差がついた。これはビーズ精製ではリードの後半部分のクオリティが急激に落ちることが原因と考えられる。実際に、図 4 に示す通り、170 サイクル付近からビーズ精製のリードはクオリティの低下が確認される。これは insert size が短いリードでは反応サイクルの後半でアダプター領域以降を解析してしまうためであり、ライブラリー精製過程で短いフラグメントが残存してしまうことに起因すると考えられる。

それぞれのリードデータの Insert size の分布を図 5 に示した。ゲル抽出のピークは

380 bp 付近であるのに対し、ビーズ精製では 160 bp 付近がピークとなった。500 サイクルや 600 サイクルの Reagent kit で解析する際は短い Insert size では無駄になる部分が多いため得られたデータ量に対し、解析に使用できるデータ量の割合が少ないという問題が生じてしまう。一方で、ビーズ精製とゲル抽出ではインサートサイズなどの違いはあるが、SNVs 解析や Assembly では同様の結果が得られることが確認された。

#### D. 考察

地衛研全国 6 支部をカバーする検査実務者 16 名からなるゲノム解析ワーキンググループを構成することで、地衛研の体制に即したマニュアルが作成できた。ワーキンググループメンバー間で意思統一を行い、ゲノム DNA のフラグメンテーション処理時間、ライブラリーのサイズセレクション手法の違いによる影響、ならびに、サーマルサイクラーの機種の違い（リッド温度設定機能の有無）について実験的な検証データを得て、マニュアルに反映した点で、他にはない、ユニークでかつ信頼性の高いマニュアルに仕上がったといえる。

リードデータの品質とデータ量確保に大きく影響するライブラリーのサイズセレクションについては、磁性ビーズによる方法とゲル切り出しによる方法を比較した。その結果、ビーズ精製は簡便な手法ではあるが、170 サイクル付近からのクオリティが大きく低下することが分かった。その原因として、インサートサイズが短いライブラリーが多く混入することによるものと考えられた。一方、ゲル切り出しの場合は、手間のかかる方法ではあるが、ライブラリーサイズはほぼ均質化しており、インサートサイズも想定通りの物が得られることから、リードデータとして無駄になる部分が少なく、ビーズ精製に比べてより多くのデータ量を確保できることを確認した。両手法での特性を理解することで、地衛研各施設での設備や使用試薬、また、解析目的に合わせた手法の選択が可能になったといえるであ

ろう。マニュアル中では、両者の比較データを appendix として提示し、フラグメント解析が可能な施設においてはビーズ精製も可能ではあるが、それ以外の施設はゲル切り出しによる方法を推奨すると明記した。

ゲノム DNA の断片化の条件として、32°C 3 分処理の妥当性を実験的に確認した。そのうえで、ヒートリッド温度設定機能を持たないサーマルサイクラーを使用する場合の対処法について検討した。断片化は反応時間 3 分でも十分進行しており、input DNA 量を 50 ng に固定した反応系において、妥当な処理時間であると思われる。ヒートリッド温度設定機能を持たないサーマルサイクラーの場合には、反応処理中リッドを開放状態に保つことで、リッド温度を 80°C に保つのと同等の結果が得られることを確認した。様々な機種を有する地方衛生研究所が、自施設の機器を用いる場合の選択肢に対する検証データといえる

#### E. 結論

ゲノム解析の 1 次処理にあたるウェットラボ操作に関するマニュアル作成に取り組んだ。地衛研全国 6 支部をカバーする検査実務者 16 名からなるゲノム解析ワーキンググループを構成することで、各施設での解析体制に関する情報の共有、ゲノム DNA の断片化処理条件やライブラリーのサイズセレクションについての実験的検証作業を経て、地衛研の体制に即したマニュアルが作成できた。

#### F. 健康危険情報

なし

#### G. 研究発表

なし

#### H. 知的財産権の出願・登録状況

##### 1. 特許取得

なし

##### 2. 実用新案登録

なし

##### 3. その他

地衛研 ゲノム DNA ライブラリー作製マニユ

# アル QIAseq FX DNA Library kit 編

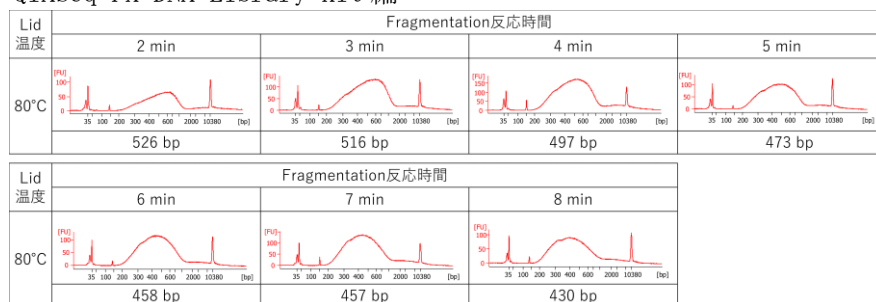


図 1. フラグメンテーションの反応時間における断片化の進行。各ライブラリの平均サイズを泳動像の下に示した。

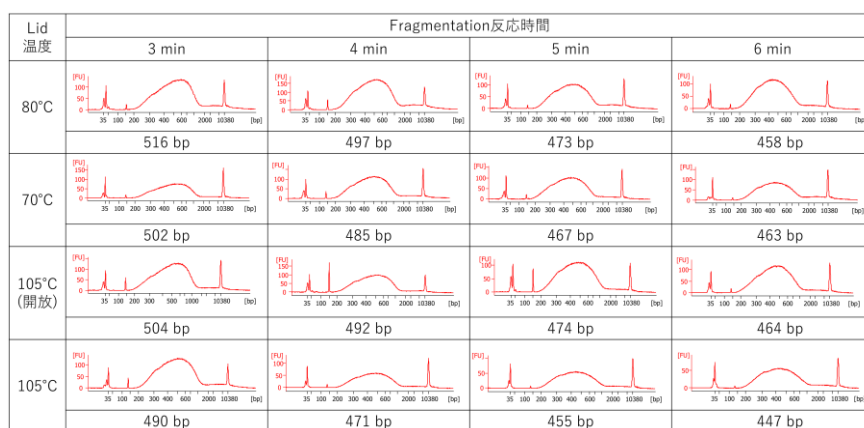


図 2. リッド温度の違いによるフラグメンテーション反応の進行。各ライブラリの平均サイズを泳動像の下に示した。

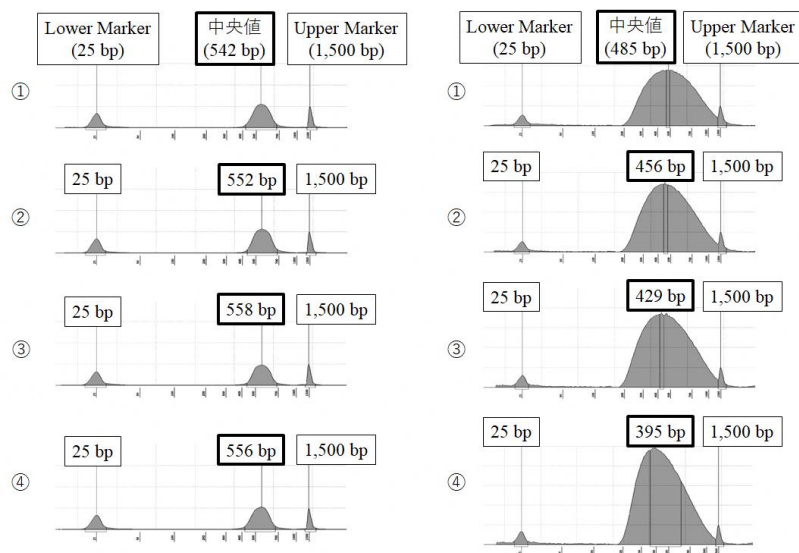


図 3. ライブラリサイズの分布。ゲル抽出 (左) とビーズ精製 (右) を TapeStation により解析した。

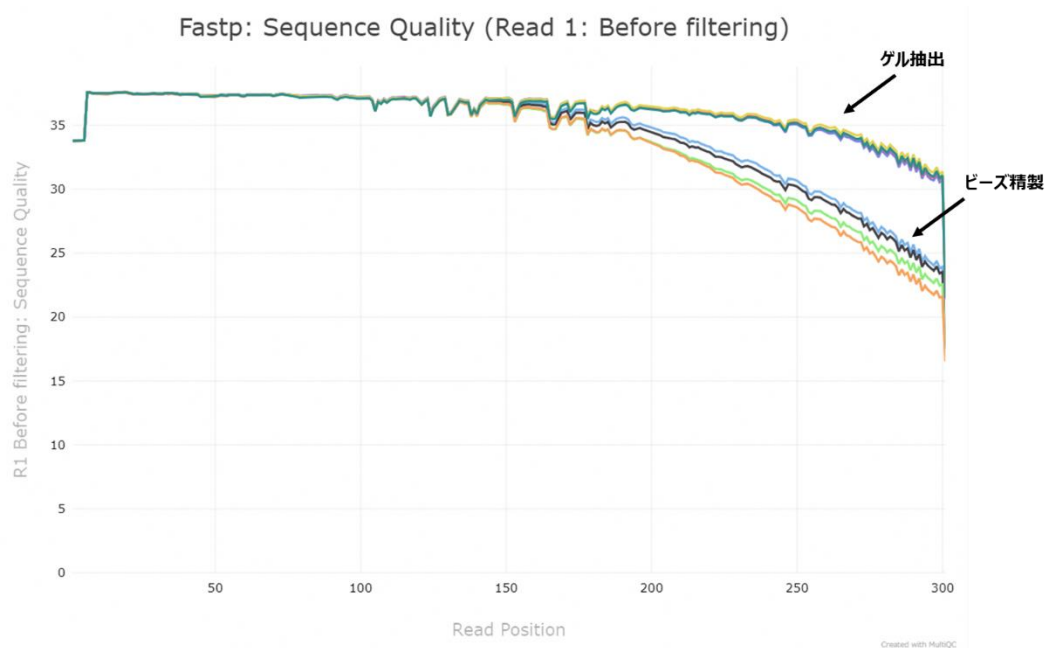


図 4. トリミング前のリードクオリティの推移。X 軸の数字は反応サイクル数と同義

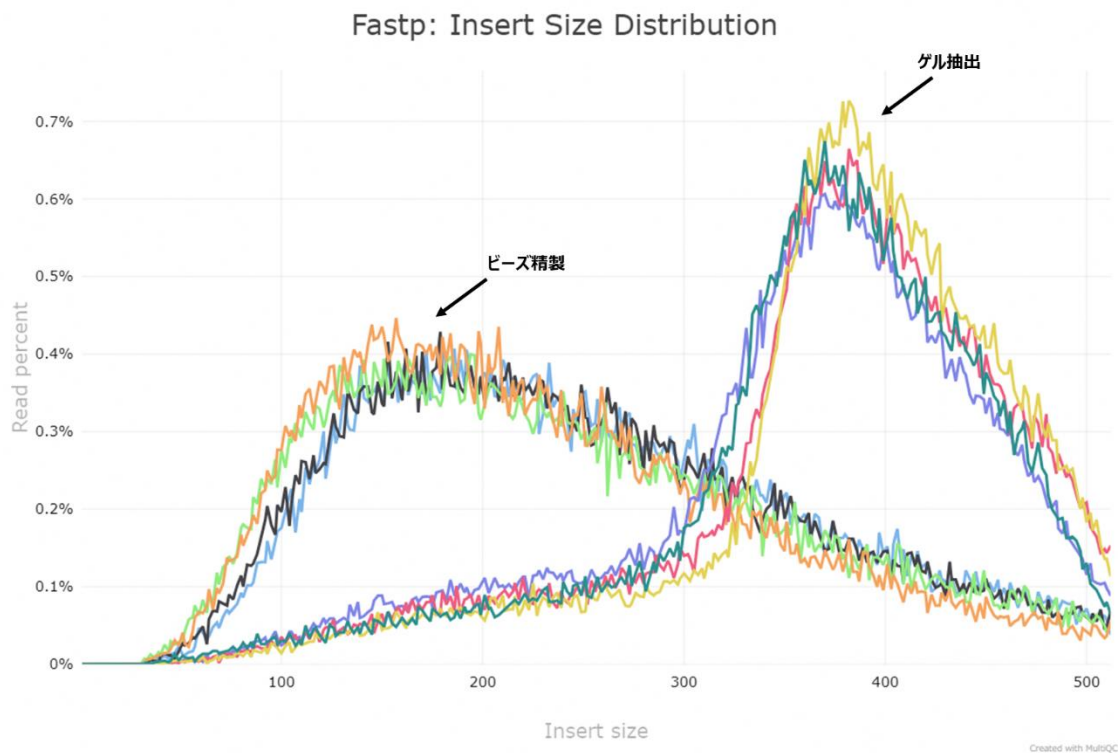


図 5. リードの Insert size の分布